5-2012

A Consensus Model for Electroencephalogram Data Via the S-Transform

Andrew Coady Young
East Tennessee State University

A Consensus Model for Electroencephalogram Data Via the S-Transform

---

A thesis

presented to

the faculty of the Department of Mathematics

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Mathematical Sciences

---

by

Andrew Young

May 2012

Jeff Knisley, Ph.D., Chair

Michele Joyner, Ph.D.

Debra Knisley, Ph.D.

Robert Gardner, Ph.D.

Keywords: Consensus Model, Signal Processing, S-Transform,

Electroencephalography, Homological Family

ABSTRACT

A Consensus Model for Electroencephalogram Data Via the S-Transform

by

Andrew Young

A consensus model combines statistical methods with signal processing to create a better picture of the family of related signals. In this thesis, we will consider 32 signals produced by a single electroencephalogram (EEG) recording session. The consensus model will be produced by using the S-Transform of the individual signals and then normalized to unit energy. A bootstrapping process is used to produce a consensus spectrum. This leads to the consensus model via the inverse S-Transform of the consensus spectrum. The method will be applied to both a control and experimental EEG to show how the results can be used in clinical settings to analyze experimental outcomes.

DEDICATION

I dedicate this thesis to Wounded Warriors. Your sacrifices are appreciated.

http://www.woundedwarriorproject.org/

ACKNOWLEDGMENTS

I would like to thank Dr. Knisley for his support and guidance through this thesis. His patience and kindness made this enjoyable. I would also like to thank my committee members who have helped me along the way, Dr. Michele Joyner, Dr. Debra Knisley, and Dr. Robert Gardner. My appreciation goes to Dr. Eric Sellers and David Ryan for motivating this project. Thanks also to ETSU for providing support and a great learning environment. My deepest appreciation goes out to my family for their love and support. And to my wife Ada, you are the best.

CONTENTS

LIST OF FIGURES

1 INTRODUCTION AND BACKGROUND

Tumors, strokes, amyotrophic lateral sclerosis, and other brain disorders have an impact on a patient's ability to function. The syndrome known as Locked-in can occur from trauma, medication overdose, strokes, and other brain damage [7]. This condition prevents an awake patient from communicating. The degree of being locked-in varies from only being able to blink the eyes to being totally locked-in where the patient has total paralysis.

Computers and related technology such as letter boards are often used to facilitate communication. However, in order for a patient to communicate through a letter board the patient must be able to acknowledge the correct letter [6]. A Locked-in patient with no motor-sensory ability still has a response which is more subtle, in the form of electrical impulses from the brain to the scalp [6]. These signals can be measured and can be used to control letter selection in a *Brain Computer Interface* (BCI).

Electroencephalography (EEG) is the study of electrical impulses from the scalp [19]. The EEG harness used in this research has 32 electrodes. Each electrode takes 256 samples per second. These 256 samples are interpolated to make a signal via a BCI [19]. A patient's response corresponds to higher amplitude areas of interest in the signal. The analysis of EEG to find the spikes in the amplitude can be approached in many ways. For instance, it is a common practice to analyze signals via the Fourier Transform [28], which is introduced later. This allows a psychologist to stack each transformed time series in the EEG family and see in the signal where the spikes in frequency are located. Currently, one approach to finding an overall pattern EEG

signals uses a moving average, where a moving average is the average of the amplitude of a signal at time n and the amplitude of the signal at the previous time, $n-1$, with the P300 BCI [19]. An alternative to a moving average is a consensus model. This technique was developed in genomics and proteomics for identifying 'hot spots' in DNA, RNA, and protein interactions [29]. In this thesis, we adapt the consensus model technology to the analysis of EEG families and show that it is superior to that of a moving average. Consensus modeling in genomics and proteomics is very successful, and we show that the same techniques are applicable to EEG analysis via the S-Transform.

1.1 Electroencephalography

Communication is vital in human interactions. A system for communication is under development for individuals whose bodies have failed yet their minds are sharp. Brain Computer Interface (BCI) is a way for humans to communicate through their brain and computer [16]. There are two types of BCI, non-invasive and invasive. Patients with ALS prefer non-invasive measures so BCI research has focused on that direction. Patients prefer EEG to invasive methods for data collection [16].

An EEG signal generates an array of signals, one signal for each node in the recording net. Figures 1 and 2 show two original signals from the N7SpllT family.

There are many techniques to analyze an array of signals. For instance, a Fourier Transformation, which we introduce later, transforms signals from the time domain to the frequency domain [28]. There is also the moving average used currently and a generalization of the Fourier Transform known as the S-Transform, which is the

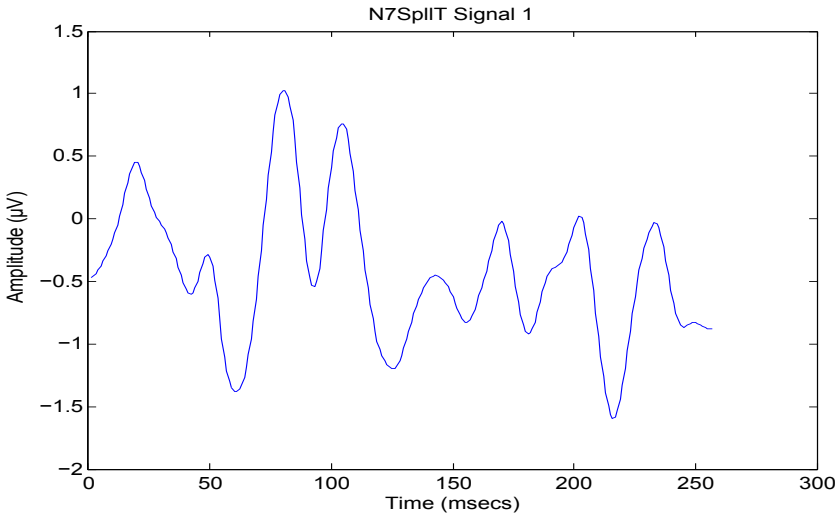

Figure 1: Signal 1 from the N7SpllT Family

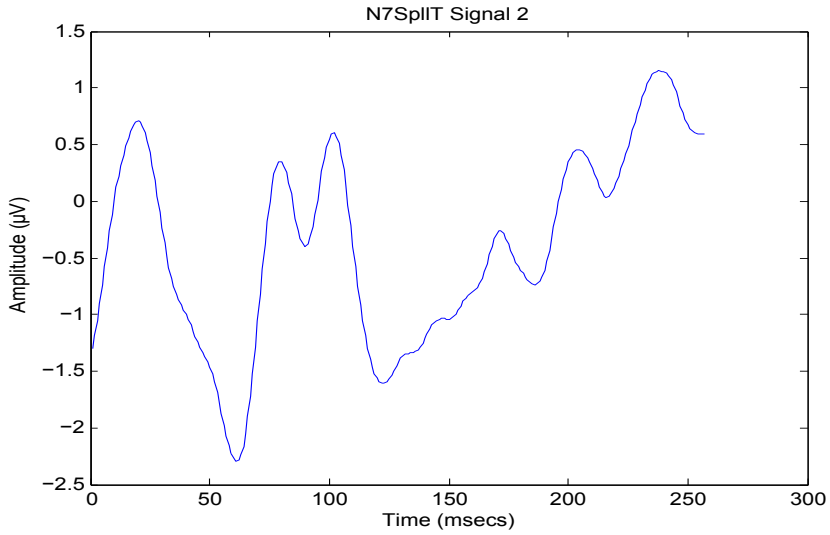

Figure 2: Signal 2 from the N7SpllT Family

focus of this project.

1.2 History Of Consensus Modeling In Genetics And Proteomics

The motivation for the use of a consensus model in EEG data is its success in genomics and proteomics. The identification of protein coding regions, which are exons in genes, is based on the presence of a period-3 characteristic that is only in specific areas of DNA [26]. The period-3 component of a DNA molecule is well known from the research of Trifonov and Sussman in 1980 where they observed the codon structure involved in the translation of base sequences into amino acids [23]. However this periodicity occurs in other coding regions which makes identifying specific coding regions challenging because these are prediction methods based on computer generated nucleotide distributions [30]. These nucleotide distribution methods are still being modified for efficiency and accuracy [30].

Each amino acid has a unique side chain which defines the chemical property of the amino acid [29]. An amino acid sequence can be converted to a sequence $\{x_k\}_{k=1}^{N}$, where each x_j is a numerical amino acid descriptor derived from side chain uniqueness. [8]. For example, Electron-Ion Interaction Potentials (EIIP) or graph theoretic invariants can be analyzed in terms of periodicity to find the characteristic frequency of a family of proteins [29]. The characteristic frequency of EIIP is determined in terms of how much of the signal is in each band of frequencies over the range of frequencies.

For example, each of the 96 PDZ domains can be represented as a sequence of EIIP. Values are analyzed in the frequency domain and the magnitude of each

frequency is recorded as a signal. Analyzing signal attributes versus the frequency is a spectral representation. Figures 3 and 4 are spectral representations of two PDZ Domains.

Taken individually, it is not clear what the characteristic frequency of PDZ domains is. In contrast, figure 5 shows the characteristic frequency of the entire family. This consensus model can subsequently be used to develop a digital filter that identifies where in a protein sequence amino acids key to the protein function can be found (i.e, the protein's 'hot spot').

1.3 Signal Spaces

A vector space consists of a nonempty set of vectors $\mathbb{V}$, a scalar field $\mathbb{F}$, vector addition in $\mathbb{V}$, and scalar multiplication between elements of $\mathbb{F}$ and $\mathbb{V}$ [9]. The particular functional spaces that are applicable to psychology are L^p spaces. An $L^p(\mathbb{R})$ space for $1 \leq p \leq \infty$ is defined to be the collection of functions f for which $|f|^p$ is Lebesgue integrable on $\mathbb{R}$ [15]. The related sequence spaces, ℓ_p, are also important. The spaces belonging to $p = 1$ and $p = 2$ are of most interest. The sequence space, ℓ^1 for $x_n \in \mathbb{R}$ is given by

$$\ell_1 = \left\{ \langle ..., x_0, x_1, ..., x_n, ... \rangle \mid \sum_{n=-\infty}^{\infty} |x_n|^1 < \infty \right\} \tag{1}$$

The space, ℓ_2, is given by

$$\ell_2 = \left\{ \langle ..., x_0, x_1, ..., x_n, ... \rangle \mid \sum_{n=-\infty}^{\infty} |x_n|^2 < \infty \right\} \tag{2}$$

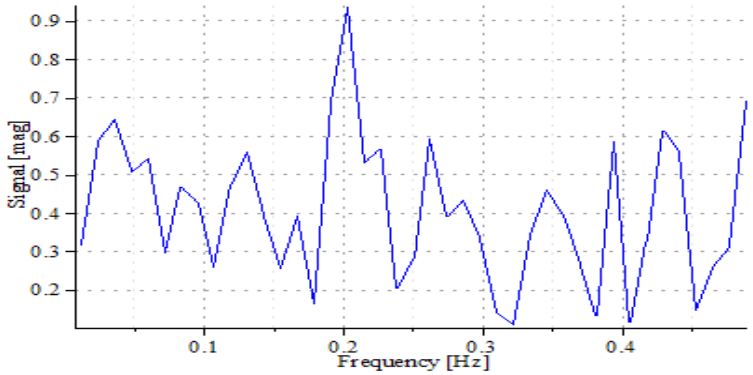

Figure 3: The Spectral Representation of a PDZ Domain

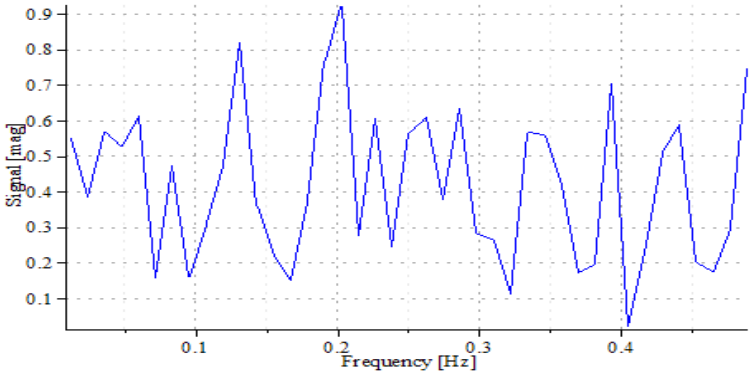

Figure 4: The Spectral Representation of a Different PDZ Domain

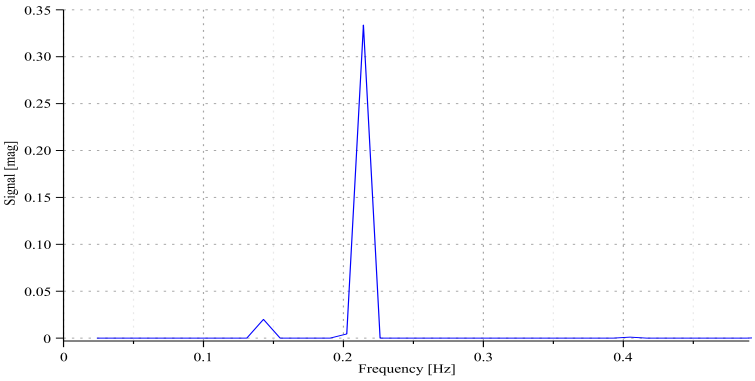

Figure 5: Consensus Spectrum of PDZ Domains

The space of absolutely integrable functions is defined

$$L^1(\mathbb{R}) = \left\{ f \mid \int_{-\infty}^{\infty} |f|^1 < \infty \right\} \tag{3}$$

where f is measurable on $\mathbb{R}$ and the integral is the Lebesgue integral. The function space L^2 is

$$L^2(\mathbb{R}) = \left\{ f \mid \int_{-\infty}^{\infty} |f|^2 < \infty \right\} \tag{4}$$

where f is measurable on $\mathbb{R}$ and the square of the Lebesgue integral is finite.

Functional analysis is a branch of mathematical analysis, the core of which is formed by the study of vector spaces endowed with some kind of limit-related structure (e.g. inner product, norm, topology, etc.) and the linear operators acting upon these spaces and respecting these structures in a suitable sense. The historical roots of functional analysis lie in the study of spaces of functions and the formulation of properties of transformations of functions such as the Fourier transform as transformations defining continuous, unitary etc. operators between function spaces. This point of view turned out to be particularly useful for the study of differential and integral equations [1].

A norm $||\cdot||$ on a linear space $\mathbb{X}$ is a mapping into $[0, \infty)$ which for $f, g \in \mathbb{X}$ (or $\mathbb{C}$ has the following properties:

- Triangle Inequality

$$||f + g|| \leq ||f|| + ||g|| \tag{5}$$

- Positive Homogeneity

$$||\alpha f|| = |\alpha| \; ||f|| \tag{6}$$

- Nonnegativity

$$||f|| \geq 0 \;\; and \;\; ||f|| = 0 \text{ if and only if } f = 0. \tag{7}$$

If $||f|| = 1$, then f is a unit function. If $||f|| \neq 1$, then $\vec{u} = \frac{f}{||f||}$ is its direction vector [15].

1.4 Signal Processing

Signal processing is an area of applied mathematics which manipulates and analyzes signals. A continuous (analog) signal is defined on all of $\mathbb{R}$ [3]. For example, a stopwatch with a second hand that has no ticks can span all the real numbers. A stopwatch that counts discretely, or equivalently via integers is an example of a discrete signal (digital). Discrete time series are digital signals. Once converted to a digital signal operations on each discrete value can be performed. A filter is a linear time invariant operator [22]. If $x, h \in \ell_1$ then their convolution $y = x * h$ is defined

$$y(n) = \sum_k h(k)x(n-k). \tag{8}$$

Filter banks generally contain a combination of two filters. High pass filters are moving differences [22]. The low frequencies of the signal are diminished or removed with a high pass filter. For example, the wavelet high pass Haar filter of a signal

$x \in \ell_1$ is given by

$$y(n) = \frac{1}{2}x(n) - \frac{1}{2}x(n-1). \tag{9}$$

Low pass filters, such as the Haar wavelet low pass filter

$$y(n) = \frac{1}{2}x(n) + \frac{1}{2}x(n-1), \tag{10}$$

are essentially moving averages. The Haar low pass averages the adjacent two components of a signal as time increases [22]. The combination (10) and (11) define the Haar Wavelet basis for $L^2(\mathbb{R})$, and similarly, wavelet filters of discrete signals are a combination of a high and lowpass filters along with upsampling and downsampling.

Multiresolution analysis uses a scaling function, $\phi(2^j t - k)$, as a basis for a set of continuous signals [22]. The wavelet, $w(2^j t - k)$, represents a wavelet basis signal at scale i and level j and time location k. The combination of the scaling function and the wavelet basis defines a multiresolution analysis that represents scale level $j+1$ as a linear combination of basis elements at scale level j. Thus, wavelets are based on scaling rather than on periodicity (frequency).

1.5 The Fourier Transform

Signal processing via the Fourier Transform assumes that continuous signals h are in the intersection $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. The Fourier Transform [21] of a signal is given by [21]

$$H(f) = \mathscr{F}[h(t)] = \int_{-\infty}^{\infty} h(t) e^{-2\pi i f t} dt \tag{11}$$

where f is the frequency domain variable in units of cycles per unit time, t. The representation at a signal as a function of frequency is called a spectral representation of the signal.

The inverse Fourier Transform is given by

$$h(t) = \mathscr{F}^{-1}[H(f)] = \int_{-\infty}^{\infty} H(f)e^{2\pi ift}df \tag{12}$$

It can be shown that if $h \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then $\mathscr{F}[h(t)]$ is also in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Thus if $h \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ then $\mathscr{F}^{-1}(\mathscr{F}(h)) = h$ [21]. For some functions the Cauchy Principal value of the Inverse Fourier transform may be necessary [12].

The Windowed or Short Time Fourier Transform (STFT) is defined as

$$STFT(\tau, f) = \int_{-\infty}^{\infty} h(t)w(t-\tau)e^{-2\pi ift}dt \tag{13}$$

where $w(t)$ is the Gaussian window,

$$w(t) = ke^{\frac{-t^2}{2}} \tag{14}$$

for k constant. The STFT incorporates a combination of a sinusoid and a localized time window. This produces time localized frequency representations.

If a continuous signal is sampled at a time interval T, then the samples form a discrete signal of length N, where N is the number of samples. The analogue of the Fourier Transform is the Discrete Fourier Transform defined

$$H\left\{n/NT\right\} = \frac{1}{N}\sum_{k=0}^{N-1} h[kT]e^{-2\pi inkT/NT} \tag{15}$$

where $k = 0...N-1$. The inverse DFT is

$$h[kT] = \sum_{n=o}^{N-1} H\left\{n/NT\right\} e^{2\pi inkT/NT}. \tag{16}$$

The limitation of the Fourier Transform is that although the spectral components of the signal can be determined; there is no information about the time at which these components exist. Moreover, the peaks in the power spectrum are independent of the phase of the signal, where a signal's phase is its state relative to a given coordinate system.

1.6 Wavelets

The Continuous Wavelet Transformation of a signal is a function,

$$W_h(\tau, a) = \int_{-\infty}^{\infty} h(t)w(t-\tau, a)dt \tag{17}$$

The Continuous Wavelet Transform (CWT) is a function of both time, t and scale a for $h \in L^2(\mathbb{R})$ and $w \in L^2(\mathbb{R})$. The Gabor Transform is

$$G_g(\tau, f) = \int_{-\infty}^{\infty} g(t)e^{\frac{-(t-\tau)^2}{2\sigma^2}} e^{-ift}dt \tag{18}$$

for $\sigma >> 0$. The Gabor kernel is

$$w(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}. \tag{19}$$

This clearly relates to the Short-Time Fourier Transform.

2 FOUNDATION OF CONSENSUS MODELING

Proteomics has established consensus modeling in the identification of protein coding regions [8]. Signal processing methods such as the Fourier Transform play a vital role in exploiting the period-3 component of DNA molecules. This foundation leads to the application of a consensus model in EEG via a variation of the Fourier transform known as the S-Transform.

2.1 Consensus Models In Proteomics

An overview of the key aspects of DNA provides a setting for this discussion. A DNA molecule has a double helix structure [26]. The four nucleotides adenine, thyomine, cytosine, and guanine denoted A, T, C, G respectively are each attached to a sugar phosphate backbone. These nucleotides pair A with T, and C with G, which are defined complementary. The upper strand of the DNA molecule is the forward genomic sequence and the bottom strand has the complementary bases. Generally one strand is active in protein synthesis.

DNA sequences can be divided into genes and intergenic spaces. Genes code proteins and are composed with four bases [26]. Sequential regions of bases in a DNA strand can be categorized into subregions called exons and introns. Protein-coding is through exons. The codon region bases can be divided into three adjacent bases that results in specific amino acids being put into place as a protein is synthesized. There are 64 possible codons. Coding sequences are determined from reading the gene left to right and defined by the codons in the exons [8]. The introns do not code proteins since they are spliced out of the RNA as part of the splicing process that results in

a mature, coding mRNA. [26]. The mRNA carries the genetic code to the ribosome which produces the protein coded by the gene.

A family of signals, sequences, residue chains where each signal is generated by a process or related system of processes is a *homological family* [27]. A protein, such as glucokinase, which can be expressed across several species, mutations to a family of proteins, and multiple site recordings such as EEG are examples of signals that, within themselves, each compose a homological family. A consensus model of a homological family is a sequence of random variables (stochastic process) whose expected value at any given time corresponds to the most likely occurring value in the homological family at that given time [20].

The long range correlation in a homological family of DNA sequences provides an opportunity to compare proteins with comparative statistical analysis of DNA. The product of the Fourier Transform of a homological family of proteins is defined as

$$P\left(n/NT\right) = \left|X_1\left(n/NT\right) X_2\left(n/NT\right) ... \; X_M\left(n/NT\right)\right|. \tag{20}$$

If a homological family of proteins has only one common function, then the power spectrum, $P(\frac{n}{NT})$ has one significant peak. The product of Fourier Transforms is defined as the consensus spectrum for the homological family of proteins. The peak frequencies are specific to each function of the protein thus termed the characteristic frequency.

2.2 Application Of Consensus Modeling In Electroencephalography

The application of consensus modeling in EEG is based on the premise that signals generated from individual electrodes represent a common phenomenon. The goal is to find the instance in time at which the electrical stimulus is maximized. However, to do so we often work in the frequency domain but working in the frequnecy domain requires a transform of the time domain signal. The Fourier transform can be used to transform EEG signals in the frequency domain.

Bootstrapping is a statistical technique in which a population is repeatedly resampled, where each resampling is a sample with replacement, and then via a bootstrap measure an empirical distribution is created [5]. A bootstrap technique can be applied to a family of EEG signals. Technology is used in bootstrapping to resample groups of signals where the probability of choosing any group of signals is the same. This estimates the sampling distribution of signals. The sampling distribution is used to make population inferences.

Bootstrapping prevents individual signals from eliminating population-wide frequencies. For example, Figure 6 contains the frequency domain representation of five signals, $\widehat{x_1}, \widehat{x_2}, \widehat{x_3}, \widehat{x_4}, \widehat{x_5}$. In four out of the five sample signals, there is an impulse at a frequency of 0.3. Without bootstrapping, as in the case with many genomic applications, the consensus spectrum is given by

$$\widehat{x_\sigma} = \prod_{i=1}^{5} \widehat{x_i}. \tag{21}$$

In our example, this implies the consensus model has a zero for a frequency of

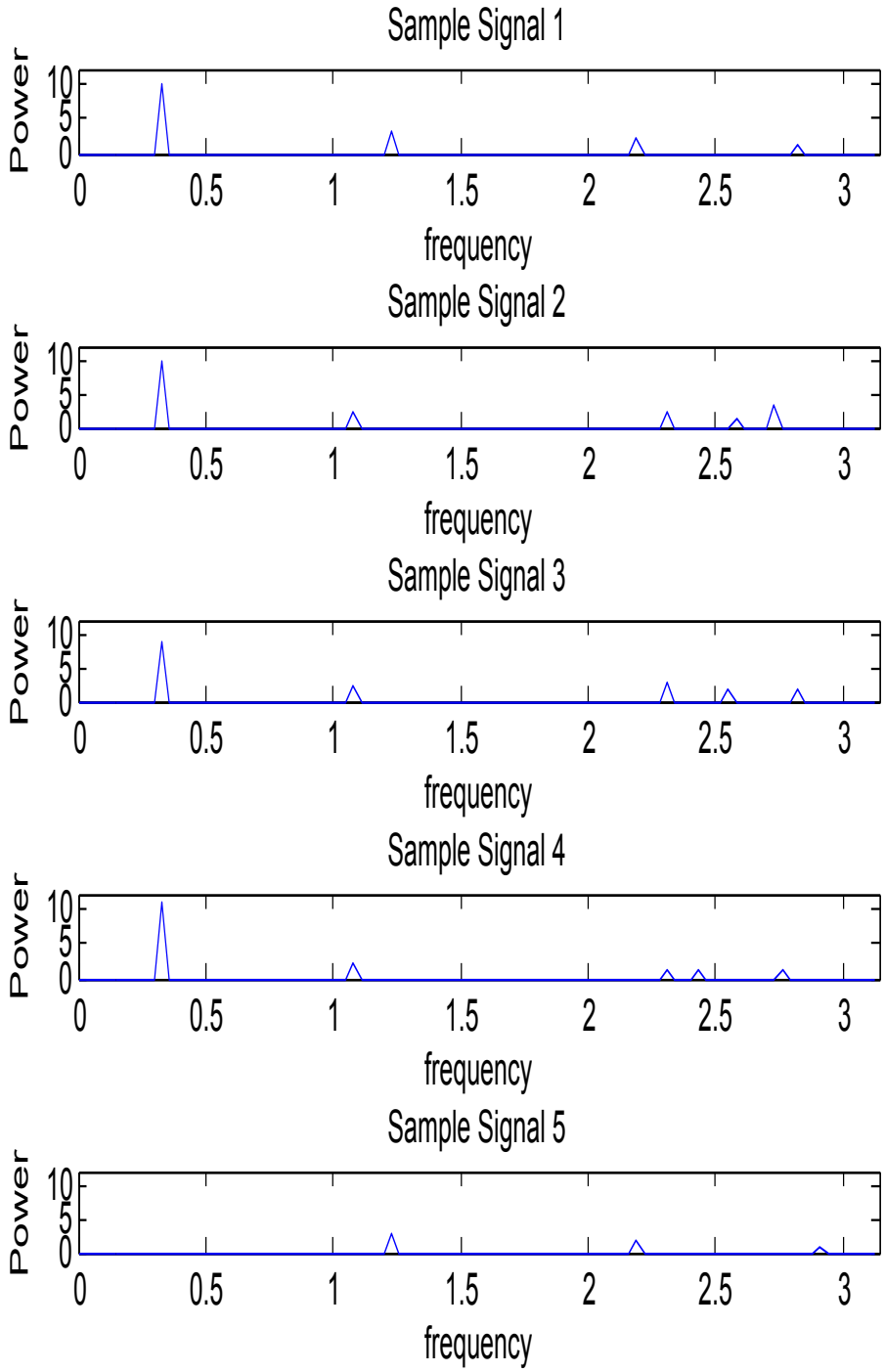

Figure 6: Sample Signals

0.3. A bootstrapping process of sampling takes n number of groups of signals and the sampling distribution would reflect that there is an impulse at a frequency of 0.3 three out of four times.

2.3 S-Transform And Its Properties

The S-Transform, developed by R.G. Stockwell [21], is defined as

$$S(\mu, a, \tau) = \int_{-\infty}^{\infty} \frac{1}{a^{\frac{1}{2}}} w\Big(\frac{t-\tau}{a}\Big) h(t) e^{-it\mu} dt \tag{22}$$

Here μ is the frequency variable in units of cycles per second, a is the dimensionless scale variable, and τ is the time location (epoch) variables in seconds. To reduce the S-Transform to a function of two variables, Stockwell defines the scale in terms of frequency as $\frac{1}{a^{\frac{1}{2}}} = |f|,\ \mu = f$. The S-Transform with scale in terms of frequency is

$$S(t, f) = \int_{-\infty}^{\infty} |f| w\Big(\big(t-\tau\big)|f|^2\big)\Big) h(t) e^{-2\pi i t f} dt \tag{23}$$

where w is a an offset Gaussian window. The S-Transform is a hybrid wavelet and STFT method that estimates the local power and phase spectrum.

The inverse S-Transform is

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} S(t, f) dt \Big) e^{2\pi i t f} df \tag{24}$$

2.4 Convolution Theorem And Its Foundational Role In Consensus Modeling

The Convolution Theorem is at the heart of Consensus Modeling because it connects multiplication in the frequency domain with autonomous (i.e. shift invariant) filters in the time domain [2]. That is, autonomous filters are defined by convolution in the time domain and by the pointwise multiplication in the frequency domain.

Theorem 2.1. *A Fourier transform of a product of signals, f and g, is the convolution of the Fourier transforms of the two signals.*

$$\mathscr{F}(f * g) = \mathscr{F}(f)\mathscr{F}(g). \tag{25}$$

Proof. This proof is reproduced from [2]. Consider $f(t)$ and $g(t)$ as signals that are elements of $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and define their inverse Fourier Transforms as:

$$f(t) = \mathscr{F}^{-1}[F(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega \tag{26}$$

$$g(t) = \mathscr{F}^{-1}[G(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} G(\omega)e^{i\omega t}d\omega. \tag{27}$$

By definition, their convolution is

$$f * g \equiv \int_{-\infty}^{\infty} g(t)f(t-\tau)dt \tag{28}$$

$$= \int_{-\infty}^{\infty} g(t)\Big[\int_{-\infty}^{\infty} F(\omega)e^{2\pi i\omega(t-\tau)}d\omega\Big]dt. \tag{29}$$

Since f, g are in $L^1(\mathbb{R})$ we can change the order of integration,

$$\begin{aligned} f * g &= \int_{-\infty}^{\infty} F(\omega)\big[g(t)e^{-2\pi i\omega t}dt\big]e^{2\pi i\omega t}d\omega \quad (30)\\ &= \int_{-\infty}^{\infty} F(\omega)G(\omega)e^{2\pi i\omega t}d\omega \quad (31)\\ &= \mathscr{F}^{-1}[F(\omega)G(\omega)](t) \quad (32) \end{aligned}$$

Since f and g are in $L^2(\mathbb{R})$ we can apply the Fourier Transform to both sides of the equation.

$$\mathscr{F}(f * g) = \mathscr{F}(f)\mathscr{F}(g) \quad (33)$$

That is, $f, g \in L^2(\mathbb{R})$ implies that $f * g$ is also in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

□

3 IMPLEMENTATION OF THE S-TRANSFORM FOR A CONSENSUS MODEL OF ELECTROENCEPHALOGRAM SIGNALS

The implementation of the S-Transform for a Consensus Model of EEG signals comes naturally with the idea of phase preservation. The goal of a consensus model in the setting of EEG analysis is to reflect the instances in the time series where there are coordinated frequencies.

3.1 Implementation

EEG signals are read into computers via a BCI in the form of a data file. To process discrete signals, we use the discrete representations of the S-Transform and Fourier Transform. The discrete S-Transform [21] of a time series $h[kT]$ is given by (letting $f \to \frac{n}{NT}$ and $\tau \to jT$),

$$S\left[jT, \frac{n}{NT}\right] = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] e^{-2\pi^2 m^2/n^2} e^{2\pi imj/N} \tag{34}$$

where $j, m, n = 0, 1, ..., N-1$. The discrete inverse S-Transform is performed by computing the Discrete Fourier Transform

$$h[kT] = \frac{1}{N} \sum_{n=0}^{N-1} \left[\sum_{j=0}^{N-1} S\left[\frac{n}{NT}, jT\right]\right] e^{2\pi ink/N} \tag{35}$$

The Matlab implementation of the S-Transform written by R.G Stockwell queries the user for a time series [21]. There are optional inputs we can use such as minimum and maximum frequency, sampling rate, and frequency sampling rate. The output

returned is a S-Transformed matrix of the input signal.

The algorithm for the input signals is a Matlab script and returns a consensus model of a collection of EEG signals. The input is a family of EEG signals in a tab-delimited format. The user enters a data file which is read into Matlab. Signals are normalized to unit energy to isolate statistical error before the transformation and bootstrapping. In the frequency domain, we use bootstrapping to produce 100 samples of eight signals randomly. This is the consensus data which we autocorrelate. The S-Transform of the consensus signal is a matrix of time and frequencies. The mesh plot and contour plot of the consensus displays the absolute amplitude at times and frequencies.

The inverse S-Transform of the autocorrelated consensus data is a single signal, the consensus model. The consensus spectrum can be used to define a convolution filter that correlates the entire family to the consensus model.

3.2 Benchmarking And Testing

The test families used for benchmarking each consists of 32 signals for a duration of 1 second sampled 256 times per second. Each signal is normalized to

$$-1 \leq \text{Amplitude} \leq 1.$$

Each test family is summarized in Appendix 2. The purpose of benchmarking and testing is to verify the consensus of families of signals matches the designed outcome. The tests of Pure Signal 1 and Pure Signal 2 have the same consensus models as

their signals since every signal in each family is the same. These tests are primarily for debugging purposes. Pure Signal 3 is a family of noisy signal frequency signals and its consensus signal is a denoised representation of the family. To test if the consensus model picks out varying chirps from 3 hertz to 7 hertz, we use mixed signals. The consensus of mixed signal 1 smoothes the collection of signals. Mixed Signal 2 produces a consensus which picks up the 3 hertz carrier and denoises the signal. Consensus Tests 1,2,3, and 4 are carrier signals with random noise. Consensus Test Consensus Test 14 is a family of signals with a 7 hertz carrier with 0.01 magnitude random noise. The consensus of this family is shown in Figure 9. In consensus test 14, the first signal of the homological family in Figure 7 and the second signal in Figure 8 is their consensus model, Figure 11.

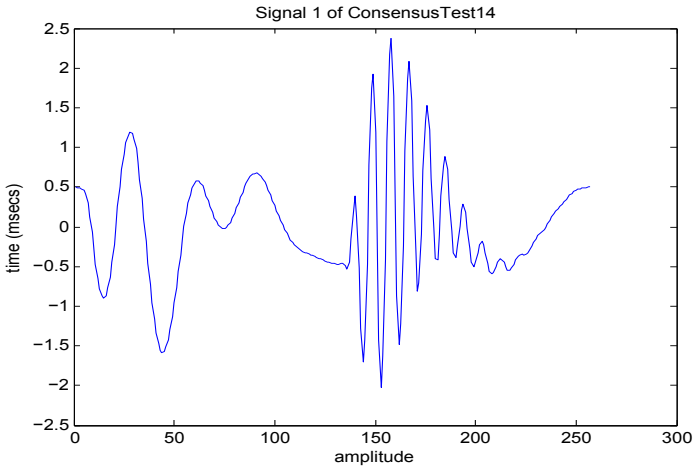

Figure 7: Signal 1 of Consensus Test 14

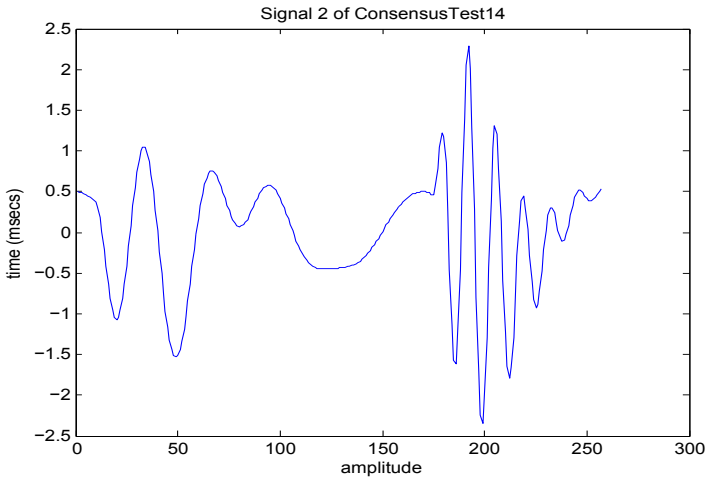

Figure 8: Signal 2 of Consensus Test 14

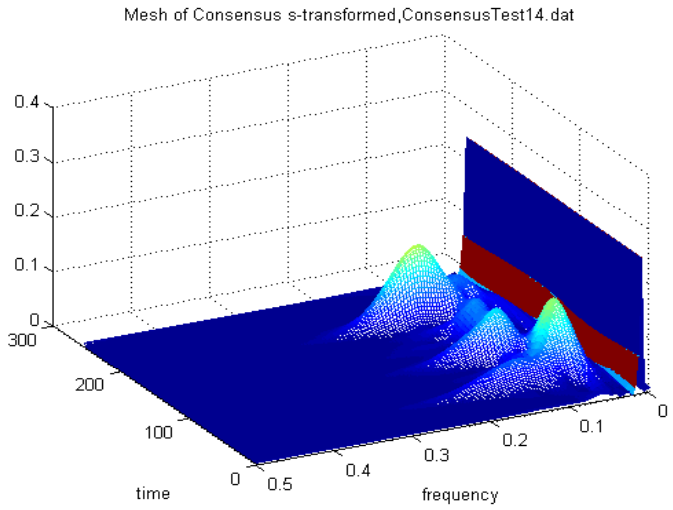

Figure 9: Mesh plot for Consensus Test 14

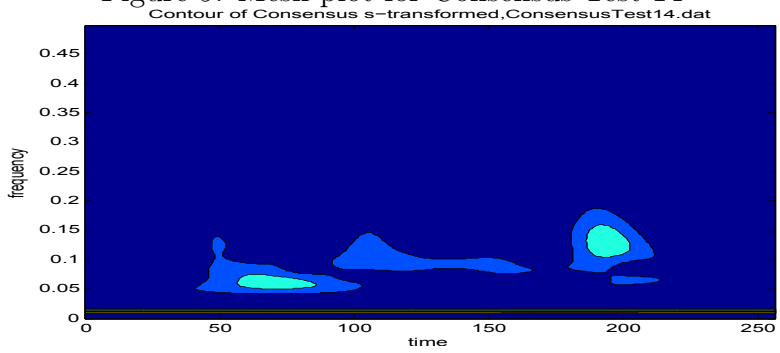

Figure 10: Contour plot for Consensus Test 14

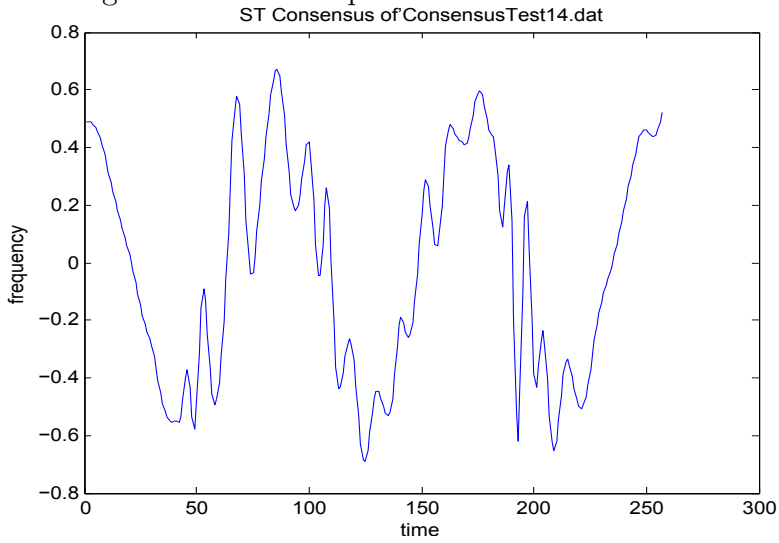

Figure 11: Consensus Signal for Consensus Test 14

3.3 Justification Of A Consensus Through Convolution

The S-Transform,

$$S_h(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-(\tau-t)^2 f^2/2} e^{-2\pi i f t} dt \tag{36}$$

can be viewed as a convolution of two functions. Reordering the terms yields

$$S_h(\tau, f) = \int_{-\infty}^{\infty} h(t) e^{-2\pi i f t} \frac{|f|}{\sqrt{2\pi}} e^{-(\tau-t)^2 f^2/2} dt. \tag{37}$$

Thus, if $A = h(t)e^{-2\pi i f t}$ and $B = \frac{|f|}{\sqrt{s\pi}} e^{-(\tau-t)^2 f^2/2}$ then

$$\begin{aligned} S_h(\tau, f) &= \int_{-\infty}^{\infty} A(t)B(\tau - t)dt & (38)\\ &= (A * B)(\tau) & (39)\\ &= \mathscr{F}^{-1}(\mathscr{F}(A) \cdot \mathscr{F}(B)) & (40)\\ &= \int_{-\infty}^{\infty} \mathscr{F}(A)(\omega)\mathscr{F}(B)(\omega)e^{2\pi\omega\tau} d\omega & (41) \end{aligned}$$

where

$$\begin{aligned} \mathscr{F}(A) &= \int_{-\infty}^{\infty} h(t)e^{-2\pi i f t} e^{-2\pi i \omega t} dt & (42)\\ &= \int_{-\infty}^{\infty} h(t)e^{-2\pi i (f+\omega) t} dt & (43)\\ &= H(\omega + f) & (44) \end{aligned}$$

where $H = \mathscr{F}(h)$ and

$$\begin{aligned} \mathscr{F}(B) &= \int_{-\infty}^{\infty} \frac{|f|}{\sqrt{2\pi}} e^{-2\pi i f t} e^{-2\pi i \omega t} dt & (45) \\ &= e^{-2\pi^2 \omega^2 / f^2}. & (46) \end{aligned}$$

Thus, the S-Transform is also given by

$$S_h(\tau, f) = \int_{-\infty}^{\infty} H(\omega + f) e^{-2\pi^2 \omega^2 / f^2} e^{2\pi i \omega} d\omega. \quad (47)$$

Suppose now that $h, k \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then

$$\begin{aligned} S_h(\tau, f) \cdot S_k(\tau, f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(\omega + f) K(\psi + f) \bullet & (48) \\ & \qquad e^{-2\pi^2 \omega^2 / f^2} e^{2\pi i \omega \tau} e^{-2\pi^2 \psi^2 / f^2} e^{2\pi i \psi \tau} d\psi d\omega & (49) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H_1(f, \omega) K_1(f, \psi) e^{2\pi i (\omega + \psi) \tau} d\psi d\omega & (50) \end{aligned}$$

where

$$\begin{aligned} H_1(f, \omega) &= e^{-2\pi^2 \omega^2 / f^2} H(\omega + f) & (51) \\ K_1(f, \psi) &= e^{-2\pi^2 \psi^2 / f^2} K(\psi + f) & (52) \end{aligned}$$

For two signals in the frequency domain, the multiplication will cancel out frequencies that are different and reinforce the frequencies that the two signals have in common.

To better illustrate, suppose that we use completing the square to write

$$
\begin{aligned}
H_1(f,\omega) &= e^{\frac{-2\pi^2\omega 2}{f^2}} \int_{-\infty}^{\infty} h(t) e^{-2\pi i\omega t} e^{-2\pi i f t} dt && (53)\\
&= \int_{-\infty}^{\infty} h(t) e^{-2\pi\left(\frac{\omega^2}{f^2}+\omega t i\right)} dt && (54)\\
&= \int_{-\infty}^{\infty} h(t) e^{\frac{-2\pi}{f^2}\left(\omega^2+\omega t i f^2-\frac{t^2 f^2}{4}+\frac{t^2 f^2}{4}\right)} e^{-2\pi i f t} dt && (55)\\
&= \int_{-\infty}^{\infty} h(t) e^{\frac{-2\pi}{f^2}\left(\omega+\frac{tif^2}{2}\right)^2} e^{\frac{-2\pi}{f^2}\left(\frac{t^2 f^4}{4}\right)} e^{-2\pi i f t} dt && (56)\\
&= \int_{-\infty}^{\infty} h(t) e^{\frac{-2\pi}{f^2}\left(\omega+\frac{tif^2}{2}\right)^2} e^{\frac{-2\pi t^2 f^2}{4}} e^{-2\pi i f t} dt && (57)\\
&= \int_{-\infty}^{\infty} h(t) e^{\frac{-2\pi t^2 f^2}{4}} e^{\frac{-2\pi}{f^2}\left(\omega+\frac{tif^2}{2}\right)^2} e^{-2\pi i f t} dt. && (58)
\end{aligned}
$$

Now there is a Gaussian modulated signal, $h(t)$. This implies that $H_1(f)K_1(f)$ will Gaussian suppress frequencies that are not in common on each $(t, t+\Delta t)$ time interval like a Short Time Fourier Transform. Specifically, $H_1(f,\omega), K_1(f,\psi)$ is an integral of 2 couples h and k only when $\omega \approx \psi$. From this it is observed that the product of two signals multiplied in the frequency domain is equivalent to the Fourier Transform of the convolution of the two signals.

$$H_1(f,\omega)K_1(f,\psi) = \mathscr{F}(h_1 * k_1) \tag{59}$$

where $h_1 = h_1(t,\omega)$ and $k_1 = k_1(t,\phi)$. Thus

$$S_h(\tau,f)S_k(\tau,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H_1(f)K_1(f)e^{2\pi i(\omega+\psi)\tau} d\psi d\omega \tag{60}$$

where $H_1(f,\omega), K_1(f,\psi)$ is a Fourier-like convolution in time (τ - variable) and also localizes in the frequency domain via the product of the Gaussians.

3.4 Consequences For Psychology

Currently the analysis of EEG sessions incorporates the Fourier Transform. The Fourier Transform can not indicate the time at which the higher frequencies occur. Instead the signals' dominant frequency overshadows the other frequencies. This implies that the higher frequency occurs for the entirety of a signal with the Fourier Transform.

The phase of the EEG Homological family is not suppressed by the S-Transform and thus a consensus model based on the S-Transform indicates the time in which the high frequencies occur. Dr. Eric Sellers of the ETSU Psychology department gave us two families of signals, N7SpllT.dat and N7SpllNT.dat. We analyzed these signals and produced consensus signals for each family in Figures 13-19. The consensus model accurately locates the time in which the higher frequencies occur. This, in turn, indicates the patient's response.

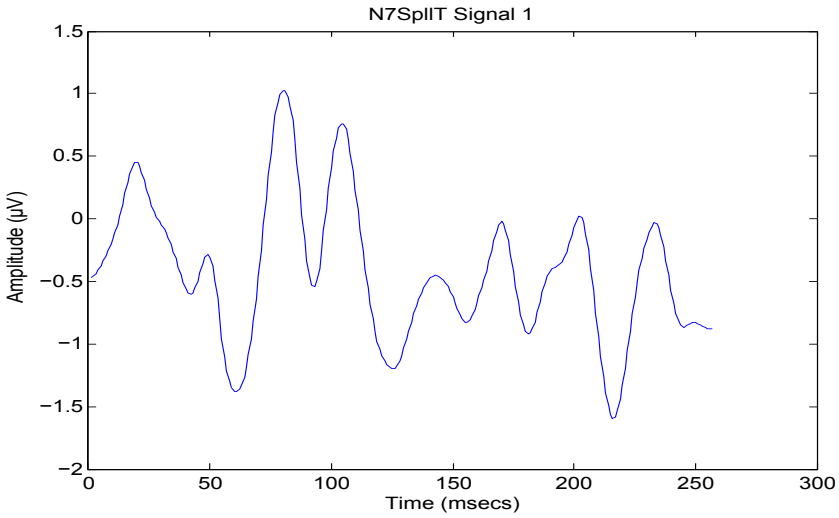

Figure 12: Signal 1 from the N7SpllT Family

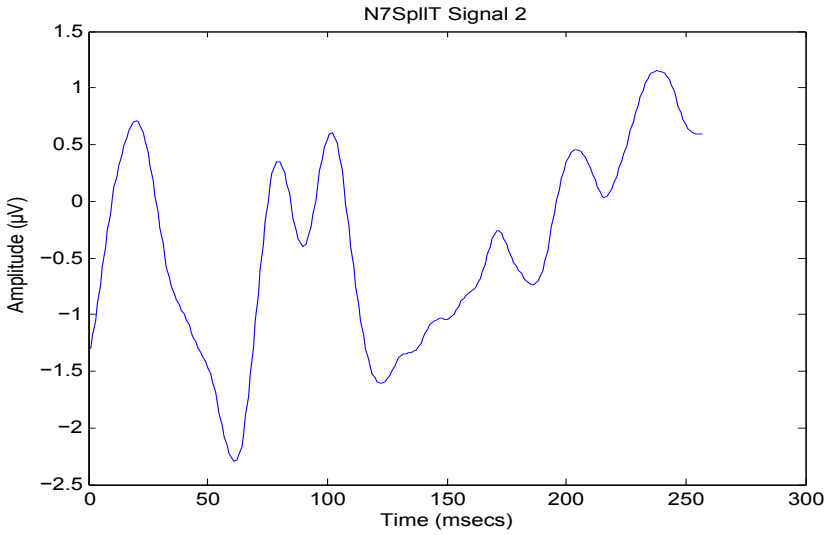

Figure 13: Signal 2 from the N7SpllT Family

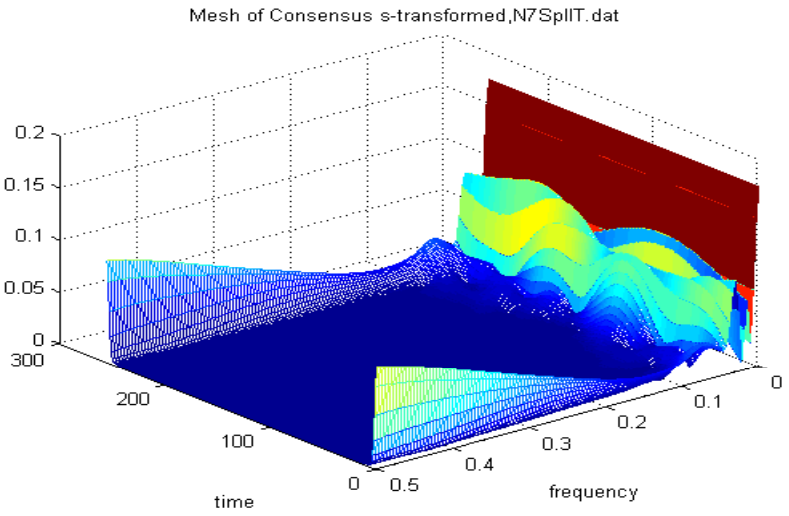

Figure 14: Mesh Plot for N7SPllT Family

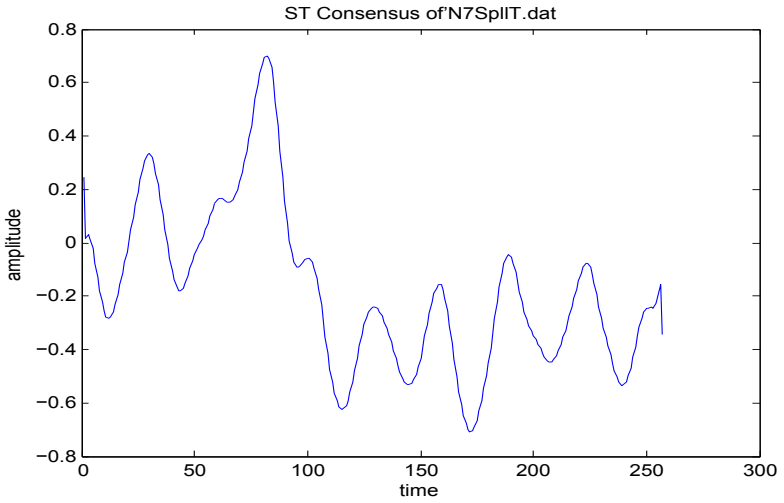

Figure 15: Consensus Signal of N7SpllT Family

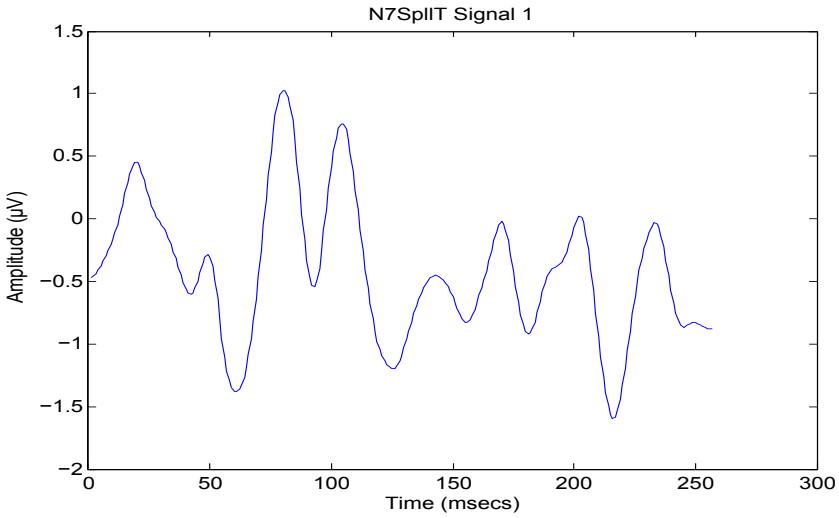

Figure 16: Signal 8 from the N7SpllNT Family

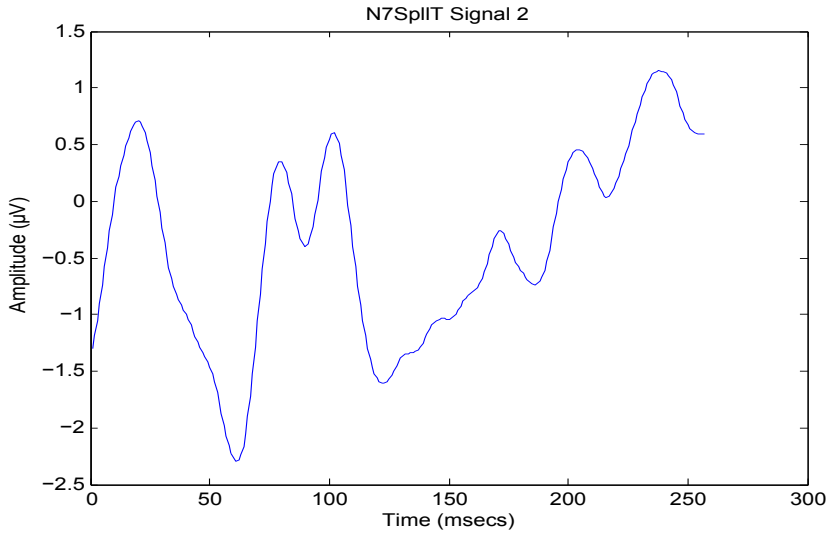

Figure 17: Signal 19 from the N7SpllNT Family

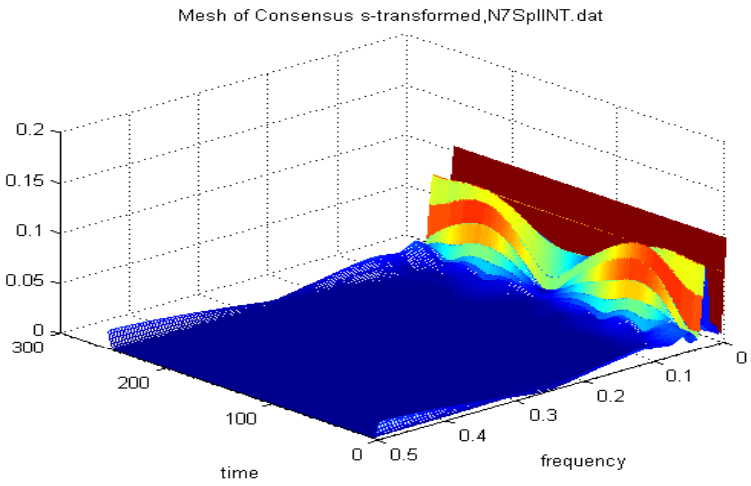

Figure 18: Mesh Plot for N7SPllNT Family

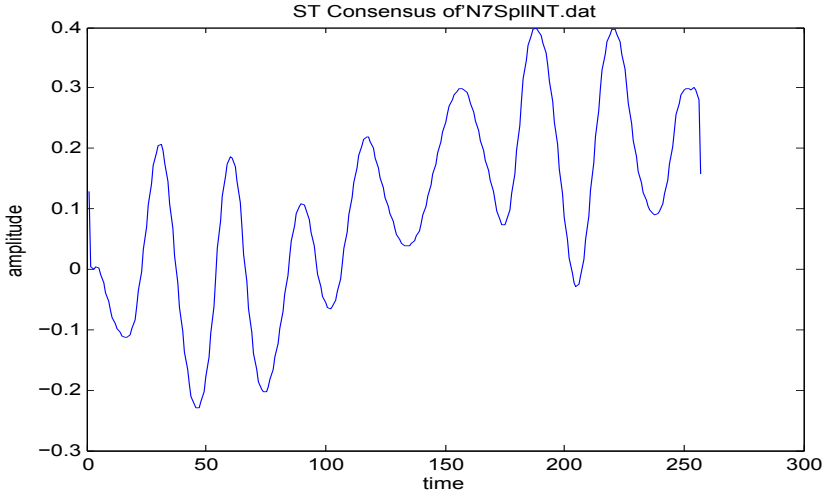

Figure 19: Consensus Signal of N7SpllNT Family

4 Conclusion And Future Directions

The techniques of consensus modeling applied to EEG data locates the time in the signal when the higher frequencies occur. This enables a BCI to locate higher frequencies of signals in a family of signals even if the magnitudes of these frequencies are not large. The ETSU Psychology department gave us two new test families to analyze, one of which have a known outcome and the other family is indeterminable as of now. We will test these families with the consensus model via the S-Transform to examine if otherwise non discript signals which are not able to be analyzed by other means can be analyzed with the S-Transform..

The next step is to share this research with the Psychology and Mathematics community through peer reviewed journal articles. Field Trip is a BCI Toolbox for Matlab. We would like to include our implementation of the S-Transform in this toolbox and other BCI systems.

BIBLIOGRAPHY

[1] C. D. Aliprantis , K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Third Edition, Springer, Berlin, Pages 462-473, 2007.

[2] A. Boggess and F. J. Narcowich, *A First Course in Wavelets with Fourier Analysis*, Prentice Hall, Texas, Pages 37-227, 2001.

[3] C. K. Chui, *An Introduction to Wavelets*, Academic Press, San Diego, Pages 49-80, 1992.

[4] J. Hanley, W. R. Adey, J. R. Zweizig, R. T. Kado, *EEG electrode-amplifier harness*, *Electroencephalography and Clinical Neurophysiology*, Ireland, Volume 30, Pages 147-150, February 1971.

[5] R. Hogg, J. McKeen, A. Craig, *Introduction to Mathematical Statistics*, 6th edition, Prentice Hall, Upper Saddle River, Page 297, 2005.

[6] D. J. Krusienski, E. W. Sellers, D. J. McFarland, T. M. Vaughan, J. R. Wolpaw. *Toward Enhanced P300 Speller Performance*, *J Neurosci Methods*, January 15, Volume 167, Pages 15-21, 2008.

[7] A. Kubler, B. Kotchoubey, J. Kaiser, J. Wolpaw, N. Birbaumer, *Brain-computer communication; Unlocking the locked in*, *Psychological Bulletin*, Vol 127(3), Pages 358-375, May 2001.

[8] S. Lin, *Digital Signal Processing for DNA Sequence Analysis*, Graduate Institute of Biomedical Electronics and Bioinformatics, Taipei, Taiwan [`http://djj.ee.ntu.edu.tw/tutorial_DNA_R98945011.pdf`], accessed April 12, 2012.

[9] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics, Raleigh, Page 160, 2000.

[10] C. Pinnegar and L. Mansinha, *The Bi-Gaussian S-Transform* , *Society for Industrial and Applied Mathematics*, online journal, volume 24, Number 5, Pages. 1678-1692, April 2003.

[11] C. Pinnegar and L. Mansinha, *The S-Transform with Windows of Arbitrary and Varying Shape* , *Geophysics*, online journal, volume 68, Number 1, January 2003.

[12] M. Plancherel, M. Leffler, *Contribution l'tude de la reprsentation d'une fonction arbitraire par les intgrales dfinies*, Rendiconti del Circolo Matematico di Palermo, Volume 30, Pages 289-335, 1910.

[13] P. Ramachandran, A. Antoniou, *Identification of Hot Spot Locations in Proteins Using Digital Filters*, *IEEE Journal of Selected Topics in Signal Processing*, Volume 2, issue 3, Pages 378-389, 2008.

[14] A. L. Rockwood, D. K. Crocket, J. R. Oliphant, and K. S. J. Elenitoba-Johnson, *Sequence Alignment by Cross-Correlation*, *Journal of Biomolecular Techniques*, Volume 16, Pages 453-458, 2005.

[15] H. L. Royden, P. M. Fitzpatrick, *Real Analysis, fourth edition*, Prentice Hall, Boston, Page 73, 2010.

[16] D. B. Ryan, *Improving Brain-Computer Interface Performance: Giving the P300 Speller Some Color*, *Masters Thesis*, East Tennessee State University, August 2011.

[17] A. Sabarish and T. Thomas, *A Frequency Domain Approach to Protein Sequence Similarity Analysis and Functional Classification*, *Signal and Image Processing: An International Journal(SIPIJ)*, Vol.2, No.1, 2011, [`http://www.scribd.com/doc/51293154/A-Frequency-Domain-Approach-to\discretionary{-}{}{}Protein-Sequence-Similarity\-Analysis-and-Functional-Classification`], accessed April 12th 2012.

[18] S. S. Sahu, G. Panda, N. V. George, *An improved S-transform for time-frequency analysis*, IEEE International Advance Computing Conference (IACC 2009) Patiala, India, Vol 6, Pages 315-319, 2009.

[19] E. W. Sellers, F. Nijboer, J. Mellinger, M. A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D. J. Krusienski, T. Vaughan, J. R. Wolpaw, N. Birbaumer, A. Kubler, *A P300-based brain-computer interface for people with amyotrophic lateral sclerosis*, *Clinical Neurophysiology*, Volume 119, Issue 8, Pages 1909-1916, August 2008.

[20] J. M. Staab, T. M. O'Connel, S. M. Gomez, *Enhancing metabolomic data anaysis with Progressive Consensus Alignment of NMR Spectra (PCANS)*

BMC Bioinformatics, [http://www.biomedcentral.com/1471-2105/11/123], accessed April 12, 2012, March 2011.

[21] R. G. Stockwell, *S-Transform*, 1996, [http://www.cora.nwra.com/~stockwel/index], accessed June 4, 2011.

[22] G. Strang and T. Nguyen, *Wavelets and Filter Banks*,Wellesley-Cambridge Press, Wellesley, Pages 1-80, 1997.

[23] E. Trifonov and J. Sussman, *The Pitch of Chromatin DNA is reflected in its Nucleotide Sequence*, Proc Natl Acad Sci USA, Volume 77, Pages 3816-3820, 1980.

[24] C. H. Trad, Q. Fang, and I. Cosic, *An Overview of Protein Sequence Comparisons Using Wavelets*, Bioelectronics Group, Department of Electrical and Computer Systems Engineering, Monash University, Clayton, Australia, 2001, [http://www.eng.monash.edu/non-cms/ecse/ieee/ieeebio2001/trad.pdf], accessed April 12, 2012.

[25] C. H. Trad, Q. Fang, and I. Cosic, *Protein Sequence Comparison Based on the Wavelet Transform Approach*, *Protein Engineering*, vol. 15, no.3, Pages 193-203, 2002.

[26] P. P. Vaidyanathan and B. Yoon, *The Role of Signal-Processing Concepts in Genomics and Proteomics*, *Journal of the Franklin Institute*, Volume 341, Pages 111-135, March 2011.

[27] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*, Wiley, West Sussux,, Pages 297-316, 2008.

[28] J. C. Woestenburg, M. N. Verbaten, J. L. Slangen, *The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain*, *Biological Psychology*, Volume 16, Issues 1-2, Pages 127-147, February-March 1983.

[29] Y. Yadav and S. Wadhwani, *Determination of Characteristic Frequency for Identification of Hot Spots in Proteins*, *International Journal of Electrical and Electronics Engineering* , Volume 1, Issue 1, Pages 26-29, 2011, [`http://interscience.in/IJEEE_vol1Iss1/paper6.pdf`], accessed April 15, 2012.

[30] C. Yin and S. Yau, *Prediction of Protein Coding Regions by the 3-Base Periodicity Analysis of a DNA Sequence*, *Journal of Theoretical Biology*, Volume 247, Pages 687-694, 2007, [`http://caw.nslij-genetics.org/gene/yin07.pdf`], accessed April 15, 2012.

APPENDICES

Appendix A: Summary of Benchmarking Signal Family

Table 1: Test Families

Pure Signal 1	32 identical signals at 7 hertz
Pure Signal 2	32 identical signals exponentially-decaying Signals at 7 hertz
Pure Signal 3	32 identical chirps at 16 hertz
Mixed Signal 1	32 identical signals varying from 3 hz to 7 hertz
Mixed Signal 2	32 identical signals with chirps, 3 hz followed by 7 hertz
Consensus Test 1	32 different signals; 3 hertz carrier with 0.01 magnitude random noise
Consensus Test 2	32 different signals; 7 hertz carrier with 0.01 magnitude random noise
Consensus Test 3	32 different signals; 3 hertz carrier with 0.1 magnitude random noise
Consensus Test 4	32 different signals; 7 hertz carrier with 0.1 magnitude random noise
Consensus Test 5	32 different signals; 3 hertz fixed with random $<$ 3 hertz signal superimposed
Consensus Test 6	32 different signals; 7 hertz fixed with random $<$ 3 hertz signal superimposed
Consensus Test 7	32 different signals; 1 hertz fixed with random $>$ 1 hertz signal superimposed
Consensus Test 8	32 different signals; amplified 3 hertz fixed with random $>$ 3 hertz signal superimposed
Consensus Test 9	32 different signals; higher amplified 3 hertz fixed with random $>$ 3 hertz signal superimposed
Consensus Test 10	32 different signals; amplified 16 hertz fixed with transient random frequency signal (5 - 30 hertz)
Consensus Test 11	32 different signals; amplified 8 hertz moving with transient random frequency signal (5 - 30 hertz)

Consensus Test 12	32 different signals; amplified 8 hertz moving with transient random frequency signal (5 - 30 hertz) + pure 3 hertz (ex 1)
Consensus Test 13	32 different signals; amplified 8 hertz moving with transient random frequency signal (5 - 30 hertz) + pure 3 hertz (ex 2)
Consensus Test 14	32 different signals; amplified 8 hertz moving with frequency based amplitude transient random frequency signal (5 - 30 hertz) + pure 3 hertz (ex 2)

Appendix B: Additional Test Images

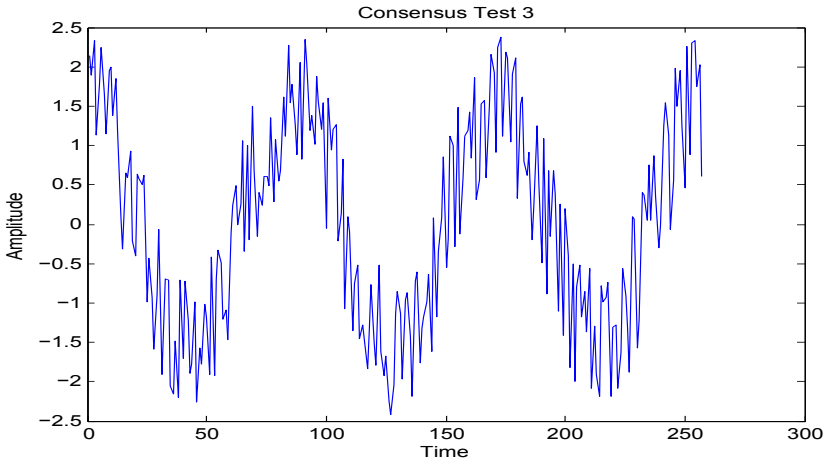

Figure 20: Signal 1 of Consensus Test 3

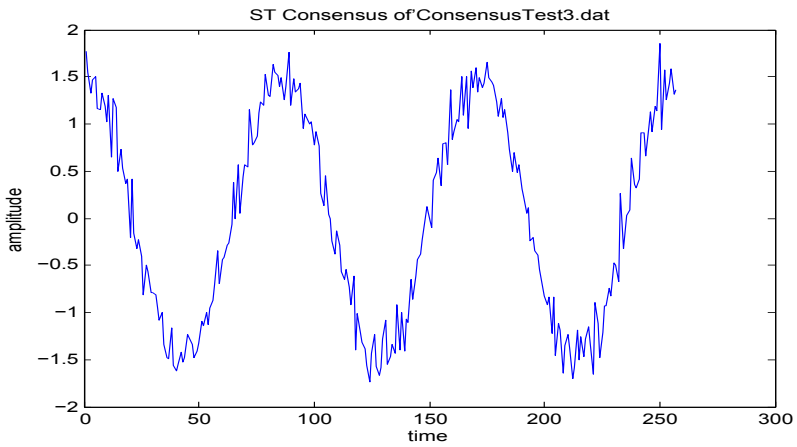

Figure 21: Consensus Signal of Consensus Test 3

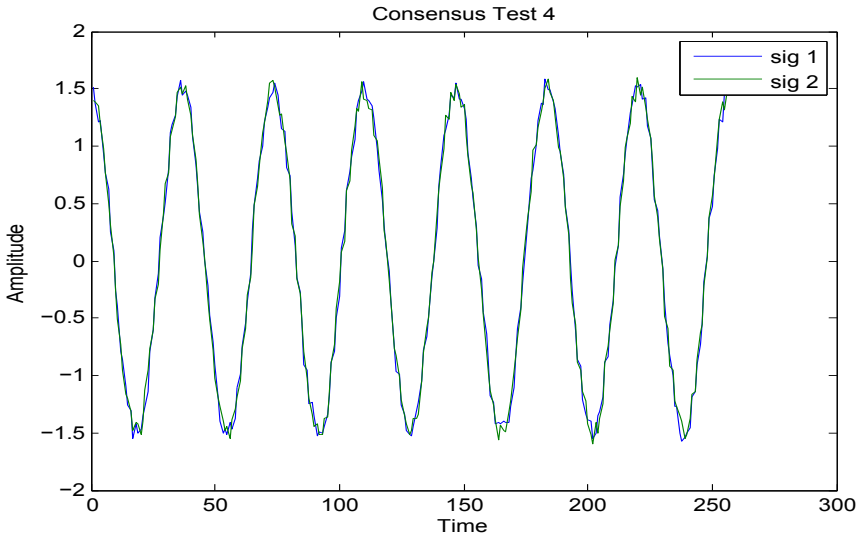

Figure 22: Signals 1 and 2 of Consensus Test 4

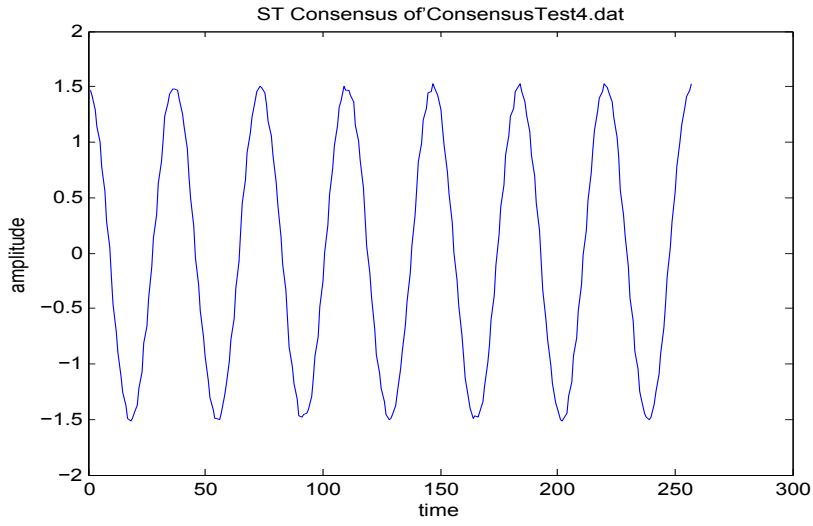

Figure 23: Consensus Signal of Consensus Test 4

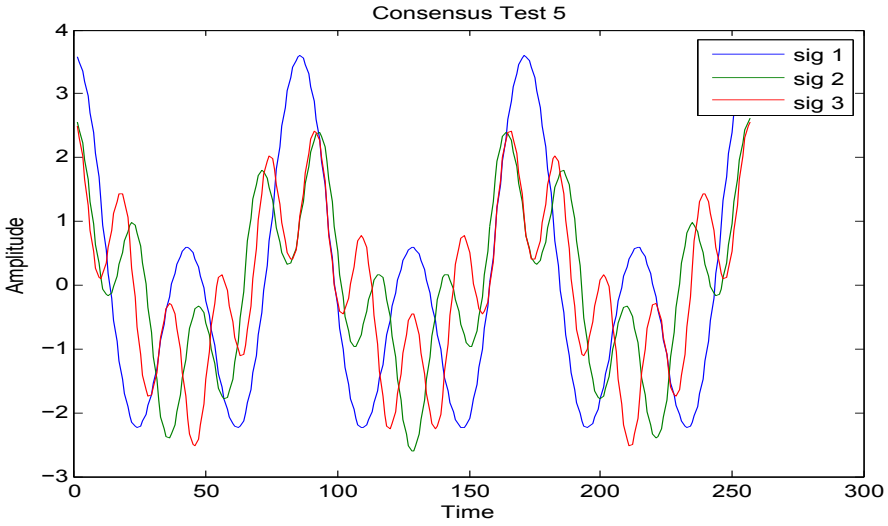

Figure 24: Signals of Consensus Test 5

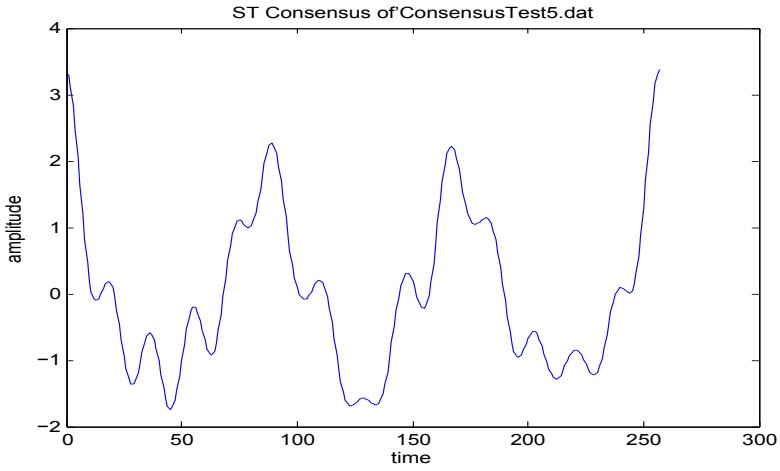

Figure 25: Consensus Signal of Consensus Test 5

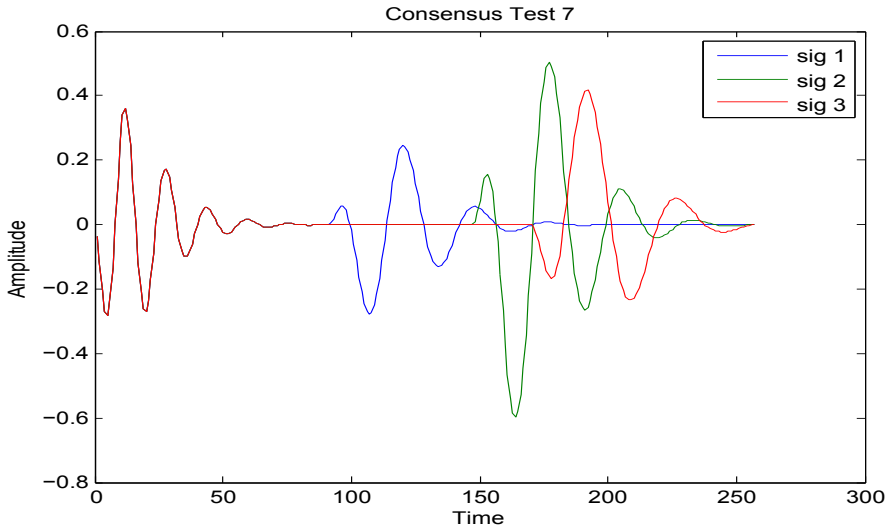

Figure 26: Signals of Consensus Test 7

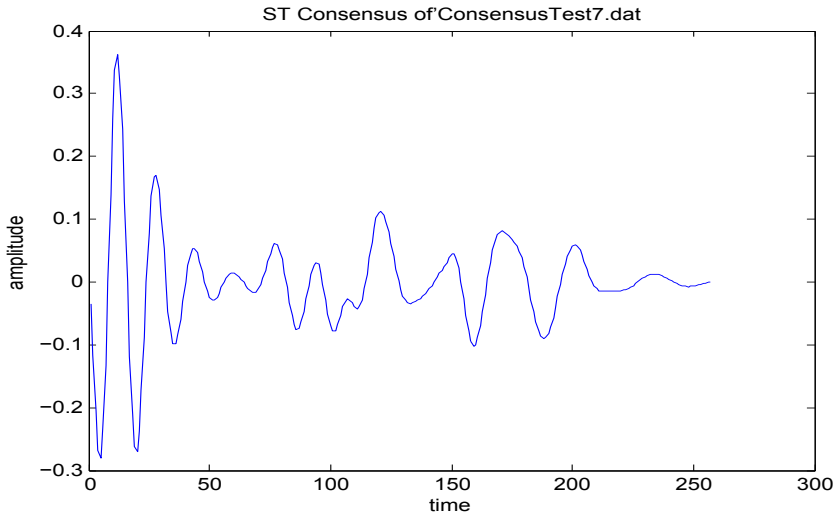

Figure 27: Consensus Signal of Consensus Test 7

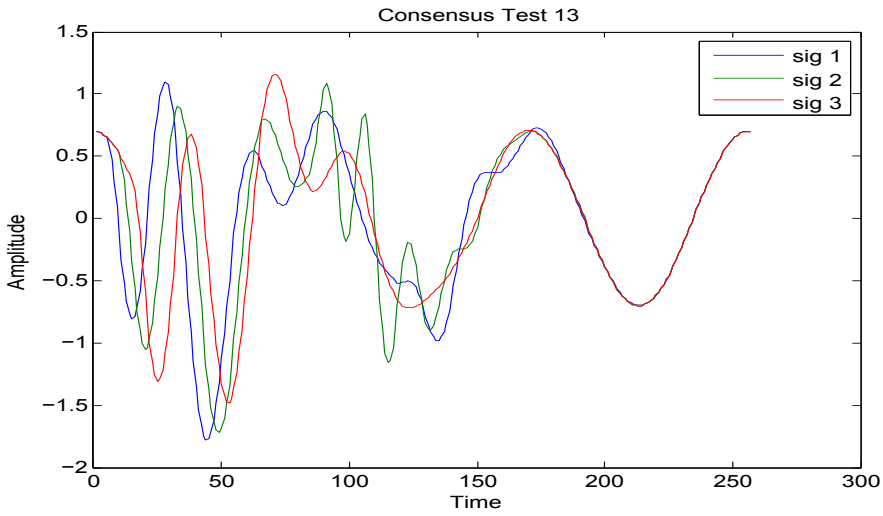

Figure 28: Signals of Consensus Test 13

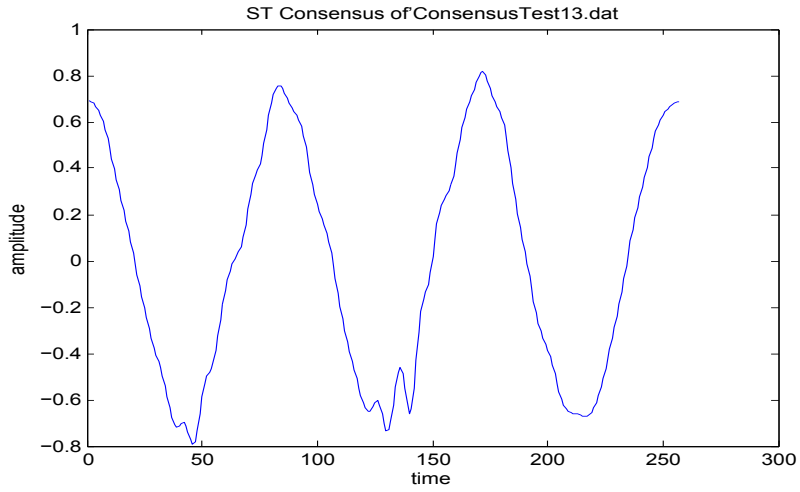

Figure 29: Consensus Signal of Consensus Test 13

Appendix C: MATLAB Code

```
%-------------------------------------------------------------------
%  steeg
%  Returns a Consensus Model of a collection of time series.
%  Code by Andrew Young
%
%--------Queued Inputs Needed-------------------------------------
%
%   a collection of timeseries in dat format for a consensus model
%   via the S-Transform
%
%--------Outputs Returned-----------------------------------------
%  figure(1)               montage of signals
%
%  figure(2)               consensus signal
%
%  figure(3)               contour plot of transformed consensus signal
%
%  figure(4)               mesh plot of transformed consensus signal
%
%  con_'filename'          array of consensus
%
%  con_spec_'filename'     array of consensus spectrum
%




%  This command prompts a user for dat file.  Once the dat file is
%  selected this routine imports the signal

[filename, pathname] = uigetfile('*.dat','Pick a dat-file')
if isequal(filename,0)| isequal(pathname,0)
    disp('User pressed cancel')
else
    disp(['User Selected ',fullfile(pathname,filename)])
end



X = load(filename);



%normalize the timeseries collection by unit energy
normbye(X);



    %S-Transform
    for i = 1:size(X,1),
    [st_matrix(ceil(size(X,2)/2)*(i-1)+1:ceil(size(X,2)/2)*i,:),
                  st_times(i,:),st_frequencies(i,:)] = st(X(i,:));
    end
    %size(st_matrix) = 4128x257
    %size(st_times) =  32 x 257
    %size(st_frequencies) = 32 x 129
```

```
    %normalize by unit energy
    matrix_normbye2(st_matrix);


   %assignment of random variables
   nrands = 100;


 %bootstrap of st_matrices
 %Here we iterate through rows and columns of st_matrix by random
 %multiples of ceil(size(X,2)/2)) which is the size(st_frequencies,2)
 %to generate a bootstrap sampling of the transformed signals.  The
 %bootstrap method takes random groups of 8 out of 32 st_matrices and
 %compose the bootstrap sample.
 for i = 1:nrands,
    for j = 1:size(X,2),
     for p = 1:8,
     randinds = randperm(32);
     pp = randinds(p);
     ConsensusData(size(st_frequencies,2)*(i-1)+1:size(st_frequencies,2)*i,j)=
          st_matrix(size(st_frequencies,2)*(pp-1)+1:size(st_frequencies,2)*pp,j);
     end
    end
 end



%multiplication of each i,j entry of bootstrap sample
%This loop takes each i,j from each sampled st_matrix and puts them into
%a vector, then each vector is multiplied through.
for f = 1:size(st_frequencies,2),
      for t = 1:size(st_times,2),
          for k = 1:size(ConsensusData,1)/size(st_frequencies,2),
              Y(k,t) = ConsensusData(size(st_frequencies,2)*(k-1)+f,t);
          end
          ConsensusData(f,t) = prod(Y(k,t),1);
      end
end

%mean of multiplied i,j entries
%In this loop each i,j entry is entered into a vector and that
%vector is averaged.
for f = 1:size(st_frequencies,2),
      for t = 1:size(st_times,2),
          for k = 1:size(ConsensusData,1)/size(st_frequencies,2),
              Y(k,t) = ConsensusData(size(st_frequencies,2)*(k-1)+f,t);
          end
          con_spec_st(f,t) = mean(Y(k,t),1);
      end
end



%inverse s-transform
Consensus = inverse_st(con_spec_st);


%plot of Consensus
 figure(2)
       plot(Consensus)
```

```
        title(strcat('ST Consensus of'' ', filename))
        xlabel('time')
        ylabel('frequency')

%s-transform of consensus
[st_matrix,st_times,st_frequencies] = st(Consensus);

%countour plot of transformed consensus
 figure(3)
   contourf(st_times,st_frequencies,abs(st_matrix));
        title(strcat('Contour of Consensus s-transformed,  ',filename))
        xlabel('time')
        ylabel('frequency')

  %mesh plot of consensus
 figure(4)
   mesh(st_times,st_frequencies,abs(st_matrix));
        title(strcat('Mesh of Consensus s-transformed,  ',filename))
        xlabel('time')
        ylabel('frequency')

 BB = abs(st_matrix);

uisave('Consensus',strcat('con_',filename));
uisave('BB',strcat('con_spec',filename));



%montage of signals
figure(1)
montage(X);
title(strcat('Montage of  ',filename))









%-------------------------------------------------------------------
%  steegk8
%  Returns a Consensus Model of Krusienski's 8.
%  (Fz,Cz,P3,Pz,P4,P07,P08,Oz)
%  Code by Andrew Young
%
%--------Queued Inputs Needed-----------------------------------------
%
%   a collection of timeseries in dat format for a consensus model
%   via the S-Transform
%
%--------Outputs Returned---------------------------------------------
%  figure(1)              montage of signals
%
%  figure(2)              consensus signal
%
%  figure(3)              contour plot of transformed consensus signal
%
```

```
%  figure(4)              mesh plot of transformed consensus signal
%
%  con_'filename'         array of consensus
%
%  con_spec_'filename'    array of consensus spectrum
%



%  This command prompts a user for dat file.  Once the dat file is
%  selected this routine imports the signal

[filename, pathname] = uigetfile('*.dat','Pick a dat-file')
if isequal(filename,0)| isequal(pathname,0)
    disp('User pressed cancel')
else
    disp(['User Selected ',fullfile(pathname,filename)])
end



X = load(filename);


%normalize the timeseries collection by unit energy
normbye(X);


%pick out the specific k8 electrodes from the collection
    k8(1,:) =  X(2,:);
    k8(2,:) = X(6,:);

        for j = 3:8,
            for i = 11:16,
                k8(j,:) = X(i,:);
            end
        end


     %S-Transform
     for i = 1:size(k8,1),
     [st_matrix(ceil(size(X,2)/2)*(i-1)+1:ceil(size(X,2)/2)*i,:),
                     st_times(i,:),st_frequencies(i,:)] = st(k8(i,:));
     end
     %size(st_matrix) = 4128x257
     %size(st_times) =  32 x 257
     %size(st_frequencies) = 32 x 129

     %normalize by unit energy
     matrix_normbye2(st_matrix);


    %assignment of random variables
    nrands = 100;


  %bootstrap of st_matrices
  %Here we iterate through rows and columns of st_matrix by random
  %multiples of ceil(size(X,2)/2)) which is the size(st_frequencies,2)
  %to generate a bootstrap sampling of the transformed signals.  The
  %bootstrap method takes random groups of 8 out of 32 st_matrices and
  %compose the bootstrap sample.
```

```
  for i = 1:nrands,
     for j = 1:size(k8,2),
      for p = 1:3,
       randinds = randperm(8);
       pp = randinds(p);
       ConsensusData(size(st_frequencies,2)*(i-1)+1:size(st_frequencies,2)*i,j)=
             st_matrix(size(st_frequencies,2)*(pp-1)+1:size(st_frequencies,2)*pp,j);
      end
     end
  end



%multiplication of each i,j entry of bootstrap sample
%This loop takes each i,j from each sampled st_matrix and puts them into
%a vector, then each vector is multiplied through.
for f = 1:size(st_frequencies,2),
      for t = 1:size(st_times,2),
          for k = 1:size(ConsensusData,1)/size(st_frequencies,2),
              Y(k,t) = ConsensusData(size(st_frequencies,2)*(k-1)+f,t);
          end
          ConsensusData(f,t) = prod(Y(k,t),1);
      end
end

%mean of multiplied i,j entries
%In this loop each i,j entry is entered into a vector and that
%vector is averaged.
for f = 1:size(st_frequencies,2),
      for t = 1:size(st_times,2),
          for k = 1:size(ConsensusData,1)/size(st_frequencies,2),
              Y(k,t) = ConsensusData(size(st_frequencies,2)*(k-1)+f,t);
          end
          con_spec_st(f,t) = mean(Y(k,t),1);
      end
end




%inverse s-transform
Consensus = inverse_st(con_spec_st);


%plot of Consensus
 figure(2)
        plot(Consensus)
        title(strcat('Consensus of  ', filename))
        xlabel('time')
        ylabel('frequency')

 %s-transform of consensus
 [st_matrix,st_times,st_frequencies] = st(Consensus);

 %countour plot of transformed consensus
  figure(3)
    contourf(st_times,st_frequencies,abs(st_matrix));
        title(strcat('Contour of Consensus s-transformed  ', filename))
        xlabel('time')
        ylabel('frequency')

   %mesh plot of consensus
```

```
  figure(4)
    mesh(st_times,st_frequencies,abs(st_matrix));
         title(strcat('Mesh of Consensus s-transformed,  ',filename))
         xlabel('time')
         ylabel('frequency')

  BB = abs(st_matrix);


 uisave('Consensus',strcat('con_',filename));
 uisave('BB',strcat('con_spec_',filename));



%montage of signals
figure(1)
montage(X);
title(strcat('Montage of    ',filename))


 function [ X ] = matrix_normbye2( X )
%matrix_normbye...normalize matrix by energy

      for i = 1:size(X,1)/ceil((size(X,2)/2)),

   energy = sum(sum(abs(X(ceil((size(X,2)/2))*(i-1)+1:129*i,:)).^2,2));
   X(ceil((size(X,2)/2))*(i-1)+1:129*i,:)=
                          X(ceil((size(X,2)/2))*(i-1)+1:129*i,:)/energy;

      end

end


function [ X ] = normbye( X )
%normbye...normalize by energy
% energy is the sum the squared terms of an array
% then divide each term of array by the energy

for i = 1:size(X,1),
energy = sum( abs( X(i,:) ).^2,2);
      X(i,:) = X(i,:)/energy;
end

end
```

VITA

ANDREW YOUNG

Education:	B.S. Mathematics, Appalachian State University, Boone, North Carolina 2009
	M.S. Mathematics, East Tennessee State University, Johnson City, Tennessee 2012
Professional Experience:	Graduate Teaching Assistant, East Tennessee State University, Johnson City, Tennessee, 2011-2012
	Graduate Assistant, Tutor, East Tennessee State University, Johnson City, Tennessee, 2010-2011
	Mathematics Teacher, Pathway Christian Academy Christiansburg, VA 2010
	Track and Field Coach, Craig County High School New Castle, Virginia 2010
	Combat Medic, U.S. Army Fort Bragg, North Carolina 2000-2006
Scholarships:	S-STEM Scholarship -research NSF sponsored for math January 2007 - December 2009
Honors Society:	Kappa Mu Epsilon